



PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/744,313A

DATE: 10/24/2002

TIME: 16:11:46

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

```

1 <110> APPLICANT: YUE, Henry
2     TANG, Y. Tom
3     AZIMZAI, Yalda
5 <120> TITLE OF INVENTION: SORTING NEXINS
7 <130> FILE REFERENCE: PF-0706 USN
9 <140> CURRENT APPLICATION NUMBER: 09/744,313A
C--> 10 <141> CURRENT FILING DATE: 2002-09-20
12 <150> PRIOR APPLICATION NUMBER: PCT/US00/14831
13 <151> PRIOR FILING DATE: 2000-05-26
15 <150> PRIOR APPLICATION NUMBER: 60/136,740
16 <151> PRIOR FILING DATE: 1999-05-27
18 <150> PRIOR APPLICATION NUMBER: 60/139,566
19 <151> PRIOR FILING DATE: 1999-06-16
21 <160> NUMBER OF SEQ ID NOS: 4
22 <170> SOFTWARE: PERL Program
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 465
26 <212> TYPE: PRT
27 <213> ORGANISM: Homo sapiens
29 <220> FEATURE:
30 <221> NAME/KEY: misc_feature
31 <223> OTHER INFORMATION: Incyte ID No: 2124842CD1
33 <400> SEQUENCE: 1
34 Met Tyr Leu Ile His Phe Cys Leu Ile Phe Arg Asn Thr Gln Lys
35   1           5           10           15
36 Arg Gly Glu Ser Phe Gly Ile Ser Arg Ile Gly Ser Lys Ile Lys
37           20           25           30
38 Gly Val Phe Lys Ser Thr Thr Met Glu Gly Ala Met Leu Pro Asn
39           35           40           45
40 Tyr Gly Val Ala Glu Gly Glu Asp Asp Phe Ile Glu Glu Gly Ile
41           50           55           60
42 Val Val Met Glu Asp Asp Ser Pro Val Glu Ala Val Ser Thr Pro
43           65           70           75
44 Asn Thr Pro Arg Asn Leu Ala Ala Trp Lys Ile Ser Ile Pro Tyr
45           80           85           90
46 Val Asp Phe Phe Glu Asp Pro Ser Ser Glu Arg Lys Glu Lys Lys
47           95          100          105
48 Glu Arg Ile Pro Val Phe Cys Ile Asp Val Glu Arg Asn Asp Arg
49          110          115          120
50 Arg Ala Val Gly His Glu Pro Glu His Trp Ser Val Tyr Arg Arg
51          125          130          135
52 Tyr Leu Glu Phe Tyr Val Leu Glu Ser Lys Leu Thr Glu Phe His
53          140          145          150

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/744,313A

DATE: 10/24/2002

TIME: 16:11:46

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

```

54 Gly Ala Phe Pro Asp Ala Gln Leu Pro Ser Lys Arg Ile Ile Gly
55                               155                               160                               165
56 Pro Lys Asn Tyr Glu Phe Leu Lys Ser Lys Arg Glu Glu Phe Gln
57                               170                               175                               180
58 Glu Tyr Leu Gln Lys Leu Leu Gln His Pro Glu Leu Ser Asn Ser
59                               185                               190                               195
60 Gln Leu Leu Ala Asp Phe Leu Ser Pro Asn Gly Gly Glu Thr Gln
61                               200                               205                               210
62 Phe Leu Asp Lys Ile Leu Pro Asp Val Asn Leu Gly Lys Ile Ile
63                               215                               220                               225
64 Lys Ser Val Pro Gly Lys Leu Met Lys Glu Lys Gly Gln His Leu
65                               230                               235                               240
66 Glu Pro Phe Ile Met Asn Phe Ile Asn Ser Cys Glu Ser Pro Lys
67                               245                               250                               255
68 Pro Lys Pro Ser Arg Pro Glu Leu Thr Ile Leu Ser Pro Thr Ser
69                               260                               265                               270
70 Glu Asn Asn Lys Lys Leu Phe Asn Asp Leu Phe Lys Asn Asn Ala
71                               275                               280                               285
72 Asn Arg Ala Glu Asn Thr Glu Arg Lys Gln Asn Gln Asn Tyr Phe
73                               290                               295                               300
74 Met Glu Val Met Thr Val Glu Gly Val Tyr Asp Tyr Leu Met Tyr
75                               305                               310                               315
76 Val Gly Arg Val Val Phe Gln Val Pro Asp Trp Leu His His Leu
77                               320                               325                               330
78 Leu Met Gly Thr Arg Ile Leu Phe Lys Asn Thr Leu Glu Met Tyr
79                               335                               340                               345
80 Thr Asp Tyr Tyr Leu Gln Cys Lys Leu Glu Gln Leu Phe Gln Glu
81                               350                               355                               360
82 His Arg Leu Val Ser Leu Ile Thr Leu Leu Arg Asp Ala Ile Phe
83                               365                               370                               375
84 Cys Glu Asn Thr Glu Pro Arg Ser Leu Gln Asp Lys Gln Lys Gly
85                               380                               385                               390
86 Ala Lys Gln Thr Phe Glu Glu Met Met Asn Tyr Ile Pro Asp Leu
87                               395                               400                               405
88 Leu Val Lys Cys Ile Gly Glu Glu Thr Lys Tyr Glu Ser Ile Arg
89                               410                               415                               420
90 Leu Leu Phe Asp Gly Leu Gln Gln Pro Val Leu Asn Lys Gln Leu
91                               425                               430                               435
92 Thr Tyr Val Leu Leu Asp Ile Val Ile Gln Glu Leu Phe Pro Glu
93                               440                               445                               450
94 Leu Asn Lys Val Gln Lys Glu Val Thr Ser Val Thr Ser Trp Met
95                               455                               460                               465
98 <210> SEQ ID NO: 2
99 <211> LENGTH: 450
100 <212> TYPE: PRT
101 <213> ORGANISM: Homo sapiens
103 <220> FEATURE:
104 <221> NAME/KEY: misc_feature
105 <223> OTHER INFORMATION: Incyte ID No: 5215690CD1

```

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/09/744,313A

TIME: 16:11:46

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

```

107 <400> SEQUENCE: 2
108 Met Glu Gln Ala Pro Pro Asp Pro Glu Arg Gln Leu Gln Pro Ala
109   1                               5                               10                               15
110 Pro Leu Glu Pro Leu Gly Ser Pro Asp Ala Gly Leu Gly Ala Ala
111                               20                               25                               30
112 Val Gly Lys Glu Ala Glu Gly Ala Gly Glu Glu Ser Ser Gly Val
113                               35                               40                               45
114 Asp Thr Met Thr His Asn Asn Phe Trp Leu Lys Lys Ile Glu Ile
115                               50                               55                               60
116 Ser Val Ser Glu Ala Glu Lys Arg Thr Gly Arg Asn Ala Met Asn
117                               65                               70                               75
118 Met Gln Glu Thr Tyr Thr Ala Tyr Leu Ile Glu Thr Arg Ser Val
119                               80                               85                               90
120 Glu His Thr Asp Gly Gln Ser Val Leu Thr Asp Ser Leu Trp Arg
121                               95                               100                              105
122 Arg Tyr Ser Glu Phe Glu Leu Leu Arg Ser Tyr Leu Leu Val Tyr
123                               110                              115                              120
124 Tyr Pro His Ile Val Val Pro Pro Leu Pro Glu Lys Arg Ala Glu
125                               125                              130                              135
126 Phe Val Trp His Lys Leu Ser Ala Asp Asn Met Asp Pro Asp Phe
127                               140                              145                              150
128 Val Glu Arg Arg Arg Ile Gly Leu Glu Asn Phe Leu Leu Arg Ile
129                               155                              160                              165
130 Ala Ser His Pro Ile Leu Cys Arg Asp Lys Ile Phe Tyr Leu Phe
131                               170                              175                              180
132 Leu Thr Gln Glu Gly Asn Trp Lys Glu Thr Val Asn Glu Thr Gly
133                               185                              190                              195
134 Phe Gln Leu Lys Ala Asp Ser Arg Leu Lys Ala Leu Asn Ala Thr
135                               200                              205                              210
136 Phe Arg Val Lys Asn Pro Asp Lys Arg Phe Thr Asp Leu Lys His
137                               215                              220                              225
138 Tyr Ser Asp Glu Leu Gln Ser Val Ile Ser His Leu Leu Arg Val
139                               230                              235                              240
140 Arg Ala Arg Val Ala Asp Arg Leu Tyr Gly Val Tyr Lys Val His
141                               245                              250                              255
142 Gly Asn Tyr Gly Arg Val Phe Ser Glu Trp Ser Ala Ile Glu Lys
143                               260                              265                              270
144 Glu Met Gly Asp Gly Leu Gln Ser Ala Gly His His Met Asp Val
145                               275                              280                              285
146 Tyr Ala Ser Ser Ile Asp Asp Ile Leu Glu Asp Glu Glu His Tyr
147                               290                              295                              300
148 Ala Asp Gln Leu Lys Glu Tyr Leu Phe Tyr Ala Glu Ala Leu Arg
149                               305                              310                              315
150 Ala Val Cys Arg Lys His Glu Leu Met Gln Tyr Asp Leu Glu Met
151                               320                              325                              330
152 Ala Ala Gln Asp Leu Ala Ser Lys Lys Gln Gln Cys Glu Glu Leu
153                               335                              340                              345
154 Val Thr Gly Thr Val Arg Thr Phe Ser Leu Lys Gly Met Thr Thr
155                               350                              355                              360

```

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/09/744,313A

TIME: 16:11:46

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

```

156 Lys Leu Phe Gly Gln Glu Thr Pro Glu Gln Arg Glu Ala Arg Ile
157          365          370          375
158 Lys Val Leu Glu Glu Gln Ile Asn Glu Gly Glu Gln Gln Leu Lys
159          380          385          390
160 Ser Lys Asn Leu Glu Gly Arg Glu Phe Val Lys Asn Ala Trp Ala
161          395          400          405
162 Asp Ile Glu Arg Phe Lys Glu Gln Lys Asn Arg Asp Leu Lys Glu
163          410          415          420
164 Ala Leu Ile Ser Tyr Ala Val Met Gln Ile Ser Met Cys Lys Lys
165          425          430          435
166 Gly Ile Gln Val Trp Thr Asn Ala Lys Glu Cys Phe Ser Lys Met
167          440          445          450
170 <210> SEQ ID NO: 3
171 <211> LENGTH: 1992
172 <212> TYPE: DNA
173 <213> ORGANISM: Homo sapiens
175 <220> FEATURE:
176 <221> NAME/KEY: misc_feature
177 <223> OTHER INFORMATION: Incyte ID No: 2124842CB1
179 <400> SEQUENCE: 3
180 gtatgaaaac tccaaaagtt atggttagcta taccttttagt ttatcatttt caaaactggt 60
181 tttctttttt tttaataatg tttttcaata aactagactg ttgtgataat ttgctatgta 120
182 ctgatacat ttttgtttaa tatcaggaa cacacagaaa aggggagaat catttggaat 180
183 cagcagaata ggtagcaaaa ttaaaggagt attcaaaagt accacaatgg agggagctat 240
184 gttgcctaata tatggtgtag ctgaagggtga agatgatttt attgaagaag gtattgttgt 300
185 aatggaagat gattctccag tggaggctgt gagcacacct aatactcccc gaaaccttgc 360
186 tgcattgaaa attagcattc catatgtaga cttttttgag gatccctcct ctgaaaggaa 420
187 ggagaaaaaa gaaagaattc ctgtgttttg tattgatgtt gaaagaaatg atagaagagc 480
188 agttggacac gagcctgaac attggtctgt ctatagaaga tatcttgaat tctatgtact 540
189 tgaatcaaaa ctaacagaat ttcatggtgc atttctgat gccagcttc cttctaagag 600
190 gatcattggc cccaaaaatt atgaattcct aaagtcaaa aggggaagagt tccaagaata 660
191 tctacagaaa cttctgcagc atccagaact gagtaatagt caacttctgg cagactttct 720
192 tttccctaata ggtggggaaa cacaatttct tgataagata ctaccagatg taaatcttgg 780
193 gaaaattata aaatctgttc ctggaaaact aatgaaagag aaaggtcagc atttggaaac 840
194 ttttatcatg aatttcatta attcttgtga gtctccaaag cctaaaccaa gtagaccaga 900
195 actgaccatt ctgagcccta cttcagaaaa caacaagaag cttttcaatg atctgtttaa 960
196 aaataatgca aaccgtgctg aaaatacaga gagaaagcaa aatcagaatt attttatgga 1020
197 ggtgatgact gtagaaggag tctatgatta cctgatgtat gtaggacggg tagttttcca 1080
198 ggttcttgac tggcttcac atctcttaat gggaactcga atcctcttta aaaacaccct 1140
199 ggaaatgtat actgattact atcttcagt taaactagaa cagctatttc aggagaccg 1200
200 tttggtctca ctcataacac ttctcagaga tgctatatc tgtgaaaaca ctgaacctcg 1260
201 ctctctccaa gataagcaaa aaggagcaaa acagactttt gaagaaatga tgaattacat 1320
202 tccagatctg ttagtcaagt gtattggtga agaaaccaag tatgaaagca tcagacttct 1380
203 gtttgatggc ttacagcaac cagtactcaa caagcagctg acttatgttt tattggacat 1440
204 ttgtatagag gaactgtttc cagagctcaa taaggtaaaa aaggaaagtt cctctgtgac 1500
205 atcttggatg taaacacttg gatttgggtat agaataaacc attgaaattt ctgctgtgcg 1560
206 aggggtggtg aaatttactt ttttgggtat attcttatat atattatgta catcgctgtc 1620
207 tgaaatttta gttatttttt gtttttaata aagactaaca caaacttaat gattaaaagt 1680
208 gattgagtct catagtcttt catttgctag ctgtgatcca aattttatta gaacataagt 1740

```

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/09/744,313A

TIME: 16:11:46

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

```

209 cacttggttat tgccattttt aaaagagaaa attcataatg atgttatggc aaacagataa 1800
210 gactgataaa ctctgtattg tatagctttg aaaataatta tgcctagtat ggagaaacag 1860
211 gaataagatc tgattttctt agagttaata tatttttagta gattggtttt cctttttttt 1920
212 attttgtaca tagttaactg tgtatctata aataaagcat cctatatgag tttttaataa 1980
213 taaaaaaaaa aa 1992
215 <210> SEQ ID NO: 4
216 <211> LENGTH: 2507
217 <212> TYPE: DNA
218 <213> ORGANISM: Homo sapiens
220 <220> FEATURE:
221 <221> NAME/KEY: misc_feature
222 <223> OTHER INFORMATION: Incyte ID No: 5215690CB1
224 <400> SEQUENCE: 4
225 cgcgccgaac tgcagccatg gagcaggcac ctccggaccc cgagcggcag ctccagccgg 60
226 cgcccttgga gccgctgggc tcccagacg ctgggctggg ggctgcggtc ggcaagggaag 120
227 cggagggggc cggagaagag agctctgggg tcgacacgat gacacacaat aatttttggg 180
228 tgaagaagat agaaatcagt gtttcagaag cagaaaaacg aactggaaga aatgccatga 240
229 acatgcaaga aacatatact gcttacctca ttgaaacaag gtcagttgaa cataccgatg 300
230 gtcagagtgt cctaacagac tcactatggc ggcgatatag tgaatttgag ttgttgagaa 360
231 gctacctttt agtttactat ccacataattg ttgtgccacc tctgccagaa aaacgggcag 420
232 aatttgtttg gcataaaactc tctgctgaca acatggatcc agattttgtg gagaggcgac 480
233 ggattggttt agaaaacttt ctcttgagga ttgcttcaca tcccatcctt tgtagagaca 540
234 aaatcttcta tctgttttta acacaggaag gtaactggaa ggagactgtg aatgaaactg 600
235 ggtttcagct gaaggcagac tccagggttaa aagcgcttaa tgcaacattc agagtgaana 660
236 acccagacaa gagatttact gaccttaagc actatagtga tgaactgcag tctgtcatct 720
237 cacatcttct tcgagtcaga gctagagtag cagatcgact ctatggtgta tataaagtac 780
238 atgggaatta tggtcgagtt ttcagtgaat ggagtgccat agaaaaagaa atgggtgatg 840
239 gactgcagag tgctggatcat catatggatg tgtatgcac tctctattgat gatattttgg 900
240 aagatgaaga acattatgca gatcagttaa aagagtatct tttttatgca gaagcattgc 960
241 gggctgtgtg caggaaacat gaacttatgc agtatgactt ggagatggct gctcaggact 1020
242 tagcatccaa gaagcagcag tgtgaggaac tggtaactgg gactgtgaga acattctctt 1080
243 tgaagggaat gactaccaag ctctttggtc aagaaactcc agagcagaga gaagccagaa 1140
244 taaaggtgct agaagaacaa ataatgaag gagaacaaca gctaaagtct aaaaatctgg 1200
245 aaggcagaga atttgtgaaa aacgcattgg ctgatattga acgcttcaaa gaacaaaaga 1260
246 accgagactt aaaggaggcc ctcataagct atgcagtcac gcagatcagt atgtgcaaaa 1320
247 aggggaattca agtttgacc aatgctaagg aatgctttag caagatgtaa tctgtgaaat 1380
248 tgaatttctc ttcaatcaaa gtgccccaaa acagaagcac aagtaaataa aagaaattta 1440
249 agtcactacc tagtatacat aaacatatac aataagttaa ataaattcag ctttctttta 1500
250 cttaattgtg gtcgtgttaa tgtagcacia aaaatatatt ttaatgaaga ttaaatatta 1560
251 taatttgagg ttttggggac tgggtgctgat tccaaaaagt taatttaata atatatacca 1620
252 acagattggt tgctacgctt ctgaaccaat gactgaatgt caagatgttc gtttaattct 1680
253 agatgtttgt ttcaagacca gctgtttcag atctattaat gtagggaatt tttccctaag 1740
254 attgaattcc tataatttact tggtaagatc cactaatctg ttatagggag ttgttttctc 1800
255 ttgccttata gttgagctat ttggtttgac aaagctcagc agaaacttga tgtgaaaaat 1860
256 ctacgatttt ctctacatt cattgatgcc ccttgtaatg ttgcataca ctagaaaatg 1920
257 ccttcaattt gtgtttttacc agaattttga tactggctcag aaattttata ctgccacaaa 1980
258 agaagactca acttctcaga tctataatgg gatacattgt catcctctag caactcctat 2040
259 atagaaagtt ttaactgaat atgttacata taagaattaa attcttctca aataattctt 2100
260 aacctcagta atgagcctaa atttactctg cttggctctc tacacatggc atttcagggt 2160

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/744,313A

DATE: 10/24/2002

TIME: 16:11:47

Input Set : A:\pf0706usn_seqlist.txt

Output Set: N:\CRF4\10242002\I744313A.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date